NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

OPTIMIZING DEPARTMENT OF DEFENSE ACQUISITION DEVELOPMENT TEST AND EVALUATION SCHEDULING

by

Shane A. Edwards

June 2015

Thesis Advisor: Gerald Brown
Second Reader: Jeffrey Kline

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704–0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2015	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE OPTIMIZING DEPARTMENT OF DEFENSE ACQUISITION DEVELOPMENT TEST AND EVALUATION SCHEDULING			**5. FUNDING NUMBERS**
6. AUTHOR(S) Shane A. Edwards			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____N/A____.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

Department of Defense (DOD) Development Test and Evaluation (DT&E) activities for new acquisitions account for a large portion of time and money during the Engineering and Manufacturing Development Phase. DOD Program Management Office test personnel develop test schedules manually using time estimates and heuristic subject matter expert advice for each test to forecast the overall time and costs associated with a developed course of action.

These manually constructed schedules take weeks to develop via many planning iterations to construct an acceptable, but not necessarily feasible or optimal solution. Ultimately, these forecast schedules and duration estimates can be inaccurate, and may result in schedule delays and/or cost overruns.

This thesis presents an optimization and simulation model as a decision support tool to improve current DT&E scheduling. We represent this resource-constrained scheduling problem as an integer linear program, and develop set enumeration reduction techniques, as well as a cascade method to reduce solve times. The proposed model, unlike current manual scheduling techniques, suggests schedules that are feasible, nearly optimal, and are produced quickly for effective analysis of alternatives.

14. SUBJECT TERMS Optimization, test and evaluation, integer linear programming, cascade, project scheduling, project management			**15. NUMBER OF PAGES** 57
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540–01-280-5500 Standard Form 298 (Rev. 2–89) Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

OPTIMIZING DEPARTMENT OF DEFENSE ACQUISITION DEVELOPMENT TEST AND EVALUATION SCHEDULING

Shane A. Edwards
Major, United States Marine Corps
B.S., United States Naval Academy, 2001

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2015

Author: Shane A. Edwards

Approved by: Gerald Brown
Thesis Advisor

Jeffrey Kline
Second Reader

Robert F. Dell
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Department of Defense (DOD) Development Test and Evaluation (DT&E) activities for new acquisitions account for a large portion of time and money during the Engineering and Manufacturing Development Phase. DOD Program Management Office test personnel develop test schedules manually using time estimates and heuristic subject matter expert advice for each test to forecast the overall time and costs associated with a developed course of action.

These manually constructed schedules take weeks to develop via many planning iterations to construct an acceptable, but not necessarily feasible or optimal solution. Ultimately, these forecast schedules and duration estimates can be inaccurate, and may result in schedule delays and/or cost overruns.

This thesis presents an optimization and simulation model as a decision support tool to improve current DT&E scheduling. We represent this resource-constrained scheduling problem as an integer linear program, and develop set enumeration reduction techniques, as well as a cascade method to reduce solve times. The proposed model, unlike current manual scheduling techniques, suggests schedules that are feasible, nearly optimal, and are produced quickly for effective analysis of alternatives.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

COA	course of action
CTP	critical technical parameter
CPM	Critical Path Method
DOD	Department of Defense
DT&E	development test and evaluation
EMD	engineering and manufacturing development
GAMS	General Algebraic Modeling System
ILP	integer linear program
MARCORSYSCOM	Marine Corps Systems Command
PERT	Program Evaluation and Review Technique
PMO	program management office
R&D	research and development
RCPSP	resource-constrained project scheduling problem
SME	subject matter expert

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Department of Defense (DOD) Development Test and Evaluation (DT&E) activities for new acquisitions account for a large portion of time and money during the Engineering and Manufacturing Development Phase. DOD Program Management Office test personnel develop DT&E schedules manually using duration estimates and heuristic subject matter expert advice for each test to forecast the duration and costs associated with a developed course of action (COA). This thesis presents an optimization and simulation model as a decision support tool to improve current DT&E scheduling.

The DT&E scheduling problem has constraints on the number and type of available test assets, the schedule of their availability, test requirements for numbers and types of these assets, and test venue constraints to accommodate test assets and perform tests. Additionally, tests have priority and precedence relationship constraints. Due to the complexity of this problem, manually constructed schedules take weeks to develop one COA and are not necessarily feasible or optimal solution. Ultimately, these forecast schedules and duration estimates can be inaccurate, and may result in schedule delays and/or cost overruns.

In this thesis, we formulate this resource-constrained scheduling problem (RSCSP) as an integer linear program (ILP) implemented in the computer program General Algebraic Modeling System (GAMS). The objective function expresses DT&E duration as well as penalties for test asset movements between venues, and penalties for any violation of precedent constraints or violations of test completions based on test priorities, a total project cost we seek to minimize.

We use test data from a previous DT&E project, developed by United States Marine Corps Systems Command (MARCORSYSCOM), to verify and validate the proposed model.

The data contains 43 tests, 36 of which are involved in partial orders (i.e., have precedence relationships). All tests are either high or medium priority, and their duration estimates are given in days. There are six available test venues located across the United States, and seven test assets available. There is only one test asset variant, and all test assets are available at the beginning of testing. The resulting ILP has 39,031 constraints and 29,271 variables, 10,590 of which are discrete variables. RSCSPs are particularly difficult to solve, so we apply set enumeration reduction techniques to reduce solve times from several hours to less than an hour. The case study results show that the differences between the MARCORSYSCOM and the GAMS model estimates are within two weeks for three-to-six month schedules.

Additionally, we use a cascade method to further speed up computation time by solving and fixing solutions for smaller ILPs over iteratively larger planning horizons. The cascade method does not necessarily suggest optimal solutions, but does provide feasible, and reasonably optimal solutions (within two weeks for three-to-six month schedules) with solve times of less than two minutes. We use the cascade method with Beta-distributed random variables for test durations to simulate overall DT&E duration, which provides a distribution of anticipated outcomes for temporal statistical analysis. The full range of simulated DT&E duration outcomes provides a better assessment of the temporal risks involved for planning and budgeting purposes than the current MARCORSYSCOM point estimates of optimistic, mean and pessimistic. This allows us to make probabilistic assessments such as probability of successfully completing all DT&E activities by a certain time, or how much time is required to complete all DT&E activities with a certain probability.

The proposed model, unlike current manual scheduling techniques, suggests schedules that are provably feasible and nearly optimal, and are produced quickly for effective analysis of alternatives.

ACKNOWLEDGMENTS

I would like to acknowledge the following individuals for their assistance.

- Distinguished Professor Gerald Brown who provided unwavering support and guidance throughout my research experience. As a leader in the optimization field, his experience and dedication was crucial to my thesis completion.

- Professor of Practice Jeffrey Kline (CAPT, USN-Ret.), who brought this project to my attention. His guidance on presentation of technical information to non-technical audiences will benefit me for the rest of my career.

- The MARCORSYSCOM engineers and test personnel, who provided me with the data that validated this model: This thesis would have been far less enlightening without their assistance and expertise..

- To COL Raymond Jones (USA-Ret.) and LtCol Mark Rafetto (USMC), who provided me with an understanding of program management and test and evaluation. Their expertise focused my research to provide capabilities that acquisition professionals need.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Department of Defense (DOD) Development Test and Evaluation (DT&E) activities for new acquisitions account for a large portion of time and money during Engineering and Manufacturing Development (EMD). DOD Instruction 5000.02 (2015) states that DT&E activities "evaluate the ability of the system to provide effective combat capability, including its ability to meet its validated and derived capability requirements" (p. 25). DOD Program Management Office (PMO) test personnel develop test schedules manually using time estimates and heuristic subject matter expert (SME) advice for each test to forecast the overall time and costs associated with a developed course of action (COA). COAs considered vary the number of test assets available, total testing time available, and which tests must be completed. These COA estimates account for variability in the individual test time estimates, but do not accurately show the variability in the overall schedule.

These manually constructed schedules also require many planning iterations to develop an acceptable, but not necessarily feasible or optimal, solution. Ultimately, these forecast schedules and duration estimates can be inaccurate, and may result in schedule delays and/or cost overruns.

This thesis introduces an optimization and simulation model as a decision support tool to suggest test schedules that minimize duration of testing, and assess different COAs. Additionally, we will show feasibility under the given constraints to determine number of test assets and/or time required to complete the test schedule.

A. CURRENT TEST SCHEDULE DEVELOPMENT

Currently, the Marine Corps Systems Command (MARCORSYSCOM), which is responsible for all major systems' acquisitions within the Marine Corps, relies on several SMEs and PMO test personnel to develop the initial test schedule for any new acquisition. This is used to establish an initial plan for coordinating test personnel, test venues, movement of test assets and a cost estimate for a program test budget.

Developing this schedule takes weeks and involves uncertainty in the length of each test due to many factors. A short list of these uncertainties includes asset maintenance problems, retesting, test facility availability conflicts, weather delays and available test personnel. Additional constraints include the number of required tests and their associated priorities, the order in which tests must be performed, the number and variants of test assets available, and the subset of test venues that can accommodate each test. These constraints often change, requiring the test personnel to rework the schedule. One objective is to minimize the length of time required to complete all testing. This typically equates to less total testing costs and enabling the acquisition program to move forward. However, the main objective is to project with accuracy the total testing time required for budgeting and coordination purposes.

1. Test Event List

Test schedule development begins with identifying the list of tests that must be conducted to evaluate given performance requirements. Each test is placed in a test functional group, known as a critical technical parameter (CTP) area, for representation on higher-level published test schedules. The Defense Acquisition Guidebook (Defense Acquisition University, 2015) states, “CTPs should focus on critical design features or risk areas (e.g., technical maturity, reliability, availability, and maintainability issues, physical characteristics or measures) that if not achieved or resolved during development will preclude delivery of required operational capabilities” (Section 9.5.3.4). For each test, the following elements are identified (see also Table 1):

- critical technical parameter (CTP) area,
- test duration estimates in days,
- asset variant type (when applicable),
- number of assets required,

- capable test venues,
- test event predecessors, and
- test event priority (Low, Medium or High).

Table 1. An excerpt of example test data that includes CTP area, precedence, venues, priority, duration, and number required.

CTP Are	Test Plan/Sheet	Prerequisite test	Test Venu	Priority of Tests	Test Days Require	Pessimist	Most Likely	Optimist	Planning Est	Multiple Vehicles Re
LM	Tilt Table	Y side slopes	ATC	H	1	3	2	1	2	No
LM	Side Slopes		ATC	H	2	7	4	3	4	No
S/HF	Initial Inspection and Safety Checkout	Y all WM test	AVTB	H	9	30	18	13	19	No
WM	Fuel Consumption - amphibious		AVTB	H	5	17	10	7	11	No
WM	Plow in testing	Y speed/powering, controlled maneuverability, max Gross vehcile	AVTB	H	2	7	4	3	4	No
S/HF	Initial Inspection and Safety Checkout	Y all test	ATC	H	9	30	18	13	19	No
S/HF	APU Noise		ATC/AVTB	M	2	7	4	3	4	No
S/HF	Climatic Chambers		ATC/YPG	M	20	67	40	29	43	No
F	Stabilization System Performance	Y Man Gun & water gunnery testing	ATC	H	3	10	6	4	6	No
F	Water Gunnery (depends on requirements)		ATC/AVTB	L	10	33	20	14	21	No
S	NBC testing		EPG	L	10	33	20	14	21	No
S	E3 testing	Y ship operations	WSMR	M	60	200	120	86	128	No

2. Asset Availability

PMO test personnel establish total assets available for testing by number and variant based on contractual delivery schedules. Individual asset availability will vary due to maintenance reliability problems that are associated with all developing technology. This is accounted for in each test duration estimate, and explained in further detail in section B of this chapter. Asset availability has a direct impact on duration of testing and is a critical element in COA analysis.

3. Time Period Availability

PMO test personnel determine the time periods available for each test and test asset. This is based on when tests must be done (i.e., seasonal climatic requirements, contractor availability, etc.), and when tests must be completed (i.e., major program milestones and/or decision points). This is done for each test asset based on previously determined asset availability.

4. Test Asset Schedule Completion

PMO test personnel use the inputs derived from the previous steps to manually construct a schedule in Microsoft Excel (Microsoft, 2011). This is meeting-intensive with multiple personnel relying on their experience to advise how to meet all the given constraints or make assumptions that effectively relax these constraints. Additional information on the computational methods used to generate these schedules will be discussed in section B of this chapter.

5. Opportunities for Improvement

The current scheduling method is time consuming, provides no statistical forecast for the duration of testing, and does not ensure optimality or feasibility (concepts to be defined more carefully later in this thesis). The optimization and simulation model developed for this thesis account for the uncertainty of each test completion time, while allowing planners to change input parameters, such as number of test assets and total test time available. The goal is to provide a model to improve planning and decision making by generating test asset schedules that can be used for analysis of alternatives and feasibility assessments.

B. PLANNING FOR UNCERTAINTY IN TEST DURATION

1. Current Deterministic Methods

The test planning horizon usually ranges from several weeks to some months, and scheduling fidelity is days. The current scheduling method uses the mean of a Beta-distribution for the calendar duration estimate of each test. The Program Evaluation and Review Technique (PERT) has been widely used since it was developed circa 1959 (Malcolm, Roseboom, Clark & Fazar, 1959), but since criticized for its assumptions (e.g., Demeulemeester & Herroelen, 2002). SMEs provide heuristic statistical parameters for the calendar duration of each test in terms of Pessimistic (P), Most Likely (M), and Optimistic (O). When these are not given, PMO test personnel generate estimates by dividing a nominal test completion time by a range of maintenance availability percentages (e.g. 0.3, 0.5 and 0.7, respectively). We can view the values of 0.3, 0.5 and 0.7 as a pessimistic case where test assets are only available for testing 30% of the time, most likely case where they are available 50% of the time, and then an optimistic case where they are available 70% of the time.

These values may change for different tests depending on perceived variability from the SMEs. The planning estimate of test duration time is then calculated by finding the mean of the Beta PERT distribution with these parameters P, M, and O, approximating for a range of six standard deviations or 99.73% of possible outcomes (Clark, 1962, p. 406). The equation for finding the mean is:

$$PlanningEstimate = \frac{P + (4 * M) + O}{6} \quad (1)$$

Because this may at first glance seem counterintuitive, we give an example of the MARCORSYSCOM estimation technique. For a test that requires seven days, the estimates for P=7/0.3=23, M=7/.05=14, and O=7/0.7=10, results in a Planning Estimate of actual calendar day test duration = 15 (see Figure 1).

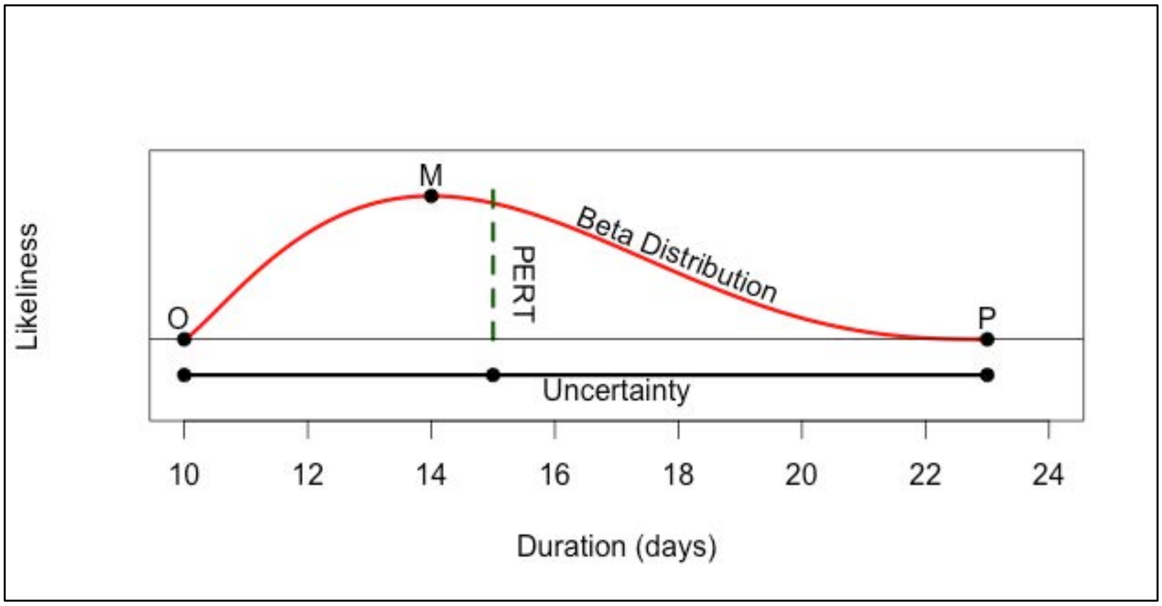

Figure 1. Visual representation of the Beta PERT distribution for the single test calendar day duration example. Pessimistic (P) duration is 23 days, Most likely (mode)(M) is 14 days, and optimistic (O) is 10 days, yielding a PERT mean of 15 days.

Each test duration planning estimate is summed up to determine the number of days required for each CTP area. The number of months required is determined by assuming there are 22 available workdays in a month (see Table 2). CTP area durations are then broken up into one-month blocks and moved around within the available time periods for each test asset in order to meet the identified test program constraints. This is also done using the pessimistic and optimistic estimates (see Table 3).

Table 2. Example of test duration planning estimates (MARCORSYSCOM) summed by CTP area to determine number of days, and then number of months required to complete all DT&E.

CTP Area	Days	Months
LM	77	3.5
F	22	1
WM	43	2
S/HF	185	8.5
Surv.	32	2
Comm	32	1
RDT/RGT	150	7
Total		**25**

Table 3. Example of a DT&E schedule produced by MARCORSYSCOM with the planning estimates provided in Table 2. CTP areas are assigned to test assets (V1 through V8) by month (1 through 9).

	P.E.	1	2	3	4	5	6	7	8	9
Planning Est.	V1 (P)	LM	LM	SH/F	SH/F	SH/F				
	V2 (P)	SH/F	LM	SH/F / LM	SH/F	SH/F				
	V3 (P)	SH/F	SH/F	F						
	V4 (P)	WM	WM	RGT						
	V5 (P)	Surv	Surv	Comm						
	V6 (P)	RDT	RDT	RGT						
	V7 (P)	RDT	RDT	RGT						
	V8 (P)	LF	LF	LF	LF	LF	LF	LF	LF	LF

Risk is then assessed by showing the three sums of minimum, planning estimate, and maximum required time to overall completion. This is a simplification of PERT that has been used since the 1950s, typically when computing power was not available to model more precisely (Malcolm et al., 1959). However, it assumes that the project's critical path remains the same, not accurately accounting for delays of their cascading effects.

2. Using Randomly-Distributed Test Durations

The model introduced in this thesis incorporates Monte Carlo simulations, randomly drawing from a Beta distribution using the same parameters for the optimistic (i.e., O, the shortest, or *min*) test duration, the pessimistic (i.e., P, the longest, or *max*) test duration, and the most likely (i.e., M, mode, $\hat{x}$) test duration. Treating these as statistical parameters for a probability density function Beta

(a,b), with a and b its parameters we can generate randomly distributed test durations following classic project scheduling advice (Malcolm et al., 1959).

$$a = 1 + 4\left(\frac{\hat{x} - min}{max - min}\right), \; b = 1 + 4\left(\frac{max - \hat{x}}{max - min}\right) \qquad \text{(2 and 3)}$$

This permits us to investigate the effects of random variations in individual test durations on completion of the entire test program by taking independent draws from a Beta distribution with $\min_t, \hat{x}_t, \text{ and } \max_t$ for each test t. This produces a histogram of outcomes from each simulated sample of project realizations, allowing for a quantitative risk assessment for that project's completion time. Additionally, this ensures that all other constraints are still accounted for, with no assumptions about the critical path.

C. ENSURING FEASIBILITY AND OPTIMIZATION

1. Ensuring Feasibility

Unlike the heuristic manual methods currently used, this model produces a Directed Acyclic Graph (DAG) representing partial order sequencing of test events and attempts to provide a minimum total completion time while honoring all incorporated constraints. Because completely feasible schedules may not always be achievable, this model will elastically penalize what appear to be necessary constraint violations, and allow such violations when necessary to meet other feasibility conditions. In this fashion, rather than declaring an ambiguous "infeasible model," all constraint violations are identified for further analysis (Brown, Dell & Wood, 1997).

2. Taking Advantage of Opportunities within Decision Variables

There are several decision variables within the network that invite efficiencies that will lead to minimizing the total test schedule time. These include which test assets are used for which tests at which venue and in what sequence. The complexity and dimensions involved in this network make it nearly impossible to select the optimal arrangement of these decision variables without the aid of optimization. This model assesses every combination for an optimal solution, seeking to minimize movement of test assets between test venues, and overall DT&E completion time.

3. Analysis of Alternatives

As previously discussed, there are several factors that can influence the input parameters for a schedule requiring continuous adjustments. These include different courses of action (COAs) being considered prior to testing (i.e., number of test assets available, total time available for testing, or projected test asset maintenance reliability). Current scheduling methods are time consuming, and limit the PMO tester’s ability to efficiently evaluate more than two or three COAs. This tool can be configured to provide increased analytics across a myriad of possible scenarios for more informed decision-making.

II. LITERATURE REVIEW

A. PROJECT PLANNING

Project planning is a problem that operations research and project management professionals have been struggling with for a long time. There have been several approaches developed using linear programming to develop schedules that optimize project completion in terms of time, cost, and/or risk. Modern (i.e., post-1950s) approaches view a project as a directed acyclic network of aggregate tasks with precedence relationships. Demeulemeester & Herroelen (2002) detail many of these, including the Program Evaluation Review Technique (PERT), the Critical Path Method (CPM), and the resource-constrained project-scheduling problem (RCPSP). These are of particular interest to this thesis as the current scheduling method uses elements of the PERT/CPM, and the test-scheduling problem presented is a RCPSP.

B. ACCOUNTING FOR UNCERTAINTY

1. Program Evaluation Review Technique and the Critical Path Method

The PERT developed by Malcolm, Roseboom, Clark and Fazar (1959) and the CPM developed by Kelley and Walker (1959) are the earliest stochastic approaches to project scheduling, and management. These methods were developed independently but have become synonymous within the project scheduling literature. Both methods focus on a “critical path” discovered within a directed network of aggregate project tasks, consisting of shortest longest path(s) through this network. A critical path is a sequence of tasks that form a longest path through the network, using point estimates for each of the task durations. The length of such a shortest longest path equates to the shortest possible project completion time. Tasks along such a path are considered critical tasks.

PERT was initially developed to probabilistically account for variability in time for completion of individual tasks, and total project completion for Navy research and development (R&D) projects. Clark (1962) explains why the Beta distribution is used to represent task completion times, and subsequently find the mean and variance of critical tasks. Additionally, Malcolm et al. (1959) assume a single critical path that tasks are independent of one another and that project duration times are normally distributed. This simplifies the statistical analysis, allowing the first two moments (mean and variance) to be summed along the critical path. Although the Central Limit Theorem (e.g., Fischer, 2011) provides strong support for such an assertion, it is not necessarily true that the distribution of durations of all PERT projects is Normal, or that this distribution will have a mean and variance equal to the sum of the means and variances on the critical path. In fact, Brown, Carlyle, Harney, Skroch, and Wood (2009) show that variability (or vulnerability) of non-critical tasks may pose substantial risks of significant project delays. These assumptions do not necessarily reflect the realities of the scheduling problem at hand, and will be explored in Chapter IV.

Fulkerson (1962) mathematically proves that PERT Beta task duration point estimates provide optimistic (i.e., shorter) project completion times compared to task durations treated as random variables. It has also been argued that the Triangle, Uniform (Elmaghraby, 1977), Gamma (Lootsma, 1966), or truncated-Weibull (Grose, 2004) distributions may be more appropriate. The Beta distribution is used in this thesis for one-to-one comparison against current MARCORSYCOM scheduling techniques.

2. Simulation

Simulation is considered the most accurate means of accounting for variability within networks, but is the most computationally expensive. Davis (2008) describes how to calculate the parameters for a Beta distribution using PERT task duration estimates. These are used to generate random values for Monte Carlo simulations, eliminating the assumptions of a single critical path, and normality (Trietsch & Baker, 2011). A similar technique is used in this thesis and is further described in Chapter IV.

Savage (2009) warns against the “flaw of averages,” which misrepresents probabilistic outcomes as a single value, typically the expected value (p.11). This is another advantage to using Monte Carlo simulations to display a probability distribution of outcomes that more accurately depict the risk involved.

C. RESOURCE-CONSTRAINED PROJECT PLANNING

Classic PERT/CPM models only consider time as the critical resource. Because we have constraints on the number and type of available test assets, the schedule of their availability, test requirements for numbers and types of these assets, and test venue constraints to accommodate test assets and perform tests in the development test and evaluation (DT&E) scheduling problem, we have a resource-constrained project-scheduling problem (RCPSP). Such problems are particularly difficult to solve. As the name implies, the RCPSP is a scheduling problem with the added complexity of resource constraints in addition to the typical temporal constraints.

Demeulemeester and Herroelen (2002) describe why RCPSPs are considered Non-Polynomial (NP)-hard, and while researchers continue to look for algorithms to solve these types of problems in polynomial time, such algorithms likely do not exist. This has led to many OR professionals to develop heuristics that reduce the problem size before searching for an optimal integer linear program (ILP) solution.

Brown, Graves and Ronen (1987) describe a cascade method for solving ILP that begins with a smaller subset of the problem, fixes that solution, and then iteratively solves larger subsets until the whole problem is solved. This significantly reduces computational time and still provides reasonably optimal solutions. A similar cascading method is implemented for this thesis and is explained in further detail in Chapter IV.

The model introduced in this thesis uses a combination of set enumeration reduction and cascading techniques to elicit PERT/CPM temporal statistical analysis from a very challenging RCPSP. Additionally, this model utilizes Monte

Carlo simulations with random task durations for additional statistical insights and comparison against classic PERT/CPM deterministic project completion times.

III. MODEL

A. PROBLEM RESTATED

There are several variants of test asset (e.g., pieces of a type of equipment to be tested) that need to be subjected to a set of test events conducted at a number of test venues (i.e., test facilities). Each test event may apply to some subset of test asset variants, and may be performed by any suitably equipped test venue.

The planning horizon consists of discrete, ordered time periods (say, days). Each test asset is to be initially delivered to a test venue at the start of a given scheduled time period, but may be subsequently moved among other venues. Completing each test event requires visiting a test venue for some given number of contiguous time periods. Moving a test asset from one test venue to another venue, and inspecting it on receipt, requires a given number of contiguous time periods. A test asset located at a test venue may be held back for other activities, and thus be unavailable for testing during some time periods. A test asset can only undergo a single test event during any time period, and each test event will be conducted at most once during the planning horizon.

Each test event has a priority (an ordered attribute), and all higher-priority test events should be started before any lower-priority ones are started, and completed before a priority-specific deadline day. Lowest-priority tests can be completed at convenience, including past the end of the planning horizon (i.e., these are optional tests).

Some tests have precedence over others, and are required to be completed before the others are started, independent of their priority. All test events of or above a given priority threshold must be completed, and the objective is to minimize completion time of the last of these tests.

Each test venue has a limit on the number of test assets it can accommodate at any time, but there is no limit on test venue capacity to perform simultaneous tests.

B. INPUTS

1. Test event data,
2. time period availability data,
3. predecessor test event sequencing,
4. test event priority (Low, Medium and High) and deadlines,
5. test asset availability data,
6. user specified asset, test event, venue and time period data,
7. distance between venues data.

C. FORMULATION

1. Index Use [~cardinality]

$t \in T$	test (alias *t'*)	[30]
$a \in A$	asset type	[3]
$a \in A_t \subseteq A$	asset types subject to test *t*	[3]
$vg \in VG$	test venue (includes element "*any*")	[7]
$v \in V \equiv VG \setminus "any"$	test venue (alias v')	[6]
$v \in V_t$	test venues capable of completing test *t*	[6]
$v \in V_{a,p}$	test venue that can receive asset *a* at start of *p*	[6]
$p \in P$	time periods in planning horizon (an ordered set) (alias p^-, p^+)	[90]
$t' \in R_t$	precedent test *t* must finish before test *t'* starts	[3]
$i \in I$	priority (an ordered set) (alias *i'*)	[3]
i_t	priority of test *t*	
$th \in TH \subseteq T$	tests that must be completed, with $i_t <\| I \|$	[28]

2. Data [units]

$a\_rec_{a,vg,p}$	type *a* assets received by venue vg at start of *p*	[assets]
$a\_type\_req_{t,a}$	number of type *a* assets required for test *t*	[assets]
$unavail_{a,v,p}$	type *a* assets unavailable at venue *v* during *p*	[assets]
a_req_t	number of assets required for test *t*	[assets]
$t\_periods_t$	periods test *t* requires	[periods]
v_cap_v	capacity of venue *v*	[assets]

$m\_periods_{v,v'}$	periods to move between facilities *v* and *v'*	[periods]
$deadline_i$	period when all priority *i* tests must be complete	[periods]
$penalty_i$	penalty for violating priority *i* deadline	[cost/period]
pri_pen	policy penalty for starting a lower priority test before a higher priority test is started	[cost]
$prec\_elastic$	indicates precedence constraints are elastic	[binary]
$prec\_pen$	penalty for not finishing precedent test before starting subsequent test	[cost]
$move\_pen$	penalty for moving assets from venue-to-venue	[cost/asset period]

3. Decision Variables [units]

Z	objective function value	[cost]
D_t	=1 if test *t* cannot be completed	[binary]
$LATE_{th}$	periods violating higher-priority test *th* deadline	[periods]
$PRI\_VIOL_{th,t'}$	=1 if a higher-priority test *th* is not started before a lower-priority test *t'* is started	[binary]
$PREC\_VIOL_{t,t'}$	=1 if a precedent test *t* is not completed before a subsequent test *t'* is started	[binary]
$W_{a,v,p}$	new test asset *a* deliveries to venue *v* at start of *p*	[assets]
$X_{t,v,p}$	=1 if test *t* begins at *v* at start of time period p	[binary]
$Y_{a,v,v',p}$	assets from facility *v* arriving at v' at start of *p* (note *v*=*v'* admitted for assets not moving)	[assets]
$S_{a,v,p}$	type *a* assets available at venue *v* during period *p*	[assets]

4. Objective Function and Constraints

$$MIN \quad Z$$

$$+(\|P\|+1)\sum_{th\in TH} D_{th} + \sum_{t\in T\setminus TH} D_t + \sum_{th\in TH} penalty_{i_{th}} LATE_{th}$$

$$+pri\_pen \sum_{th,\in TH,t'\in T|i_t>i_{t'}} PRI\_VIOL_{th,t'} + prec\_pen \sum_{t\in T|t'\in R_t} PREC\_VIOL_{t,t'}$$

$$+move\_pen \sum_{a,v,v'\neq v,p} m\_periods_{v,v'} Y_{a,v,v',p} \qquad [A0]$$

$$s.t. \quad Z \geq (p + t\_periods_{th} - 1) X_{th,v,p} \qquad \forall th \in TH, v \in V_{th}, \; p + t\_periods_{th} - 1 \leq \|P\| \qquad [A1]$$

$$\sum_{v\in V_t, p+t\_periods_t-1\leq\|P\|} X_{t,v,p} + D_t = 1 \qquad \forall t \in T \qquad [A2]$$

$$\sum_{v\in V_{th}, p\in P} (p + t\_periods_{th} - 1) X_{th,v,p} \leq deadline_{i_{th}} + LATE_{th} \qquad \forall th \in TH \qquad [A3]$$

$$\sum_{v\in V_{th}, p^- < \min\{p, \|P\| - t\_periods_{th}\}} X_{th,v,p^-} + PRI\_VIOL_{th,t'} \geq \sum_{v\in V_{t'}} X_{t',v,p} \qquad \forall th \in TH, t' \in T \mid i_{th} > i_{t'}, \; p + t\_periods_{t'} - 1 \leq \|P\| \qquad [A4]$$

$$\sum_{v'\in V_{t'}, p^- + t\_periods_{t'} + m\_periods_{v',v}|_{v\neq v'} \leq p} X_{t',v',p^-} + PREC\_VIOL_{t',t}\,|_{prec\_elastic} \geq X_{t,v,p} \qquad \forall t' \in T \mid t \in R_{t'}, v \in V_t, \; p + t\_periods_t - 1 \leq \|P\| \qquad [A5]$$

$$\sum_{v'\in V_{t'}, p^- + t\_periods_t + m\_periods_{v',v}|_{v\neq v'} = p} X_{t',v',p^-} \leq \sum_{p\leq p^+} X_{t,v,p^+} + PREC\_VIOL_{t',t}\,|_{prec\_elastic} \qquad \forall t' \in T \mid t \in R_{t'}, v \in V_t, \; p + t\_periods_t - 1 \leq \|P\| \qquad [A6]$$

$$\sum_{v\in V_{a,p}} W_{a,v,p} = a\_rec_{a,\text{"any"},p} \qquad \forall a \in A, p \in P \qquad [A7]$$

$$\sum_{v'\in V|p>m\_periods_{v',v}} Y_{a,v',v,p} + W_{a,v,p} + a\_rec_{a,v,p} \geq \sum_{v'\in V} Y_{a,v,v',p+m\_periods_{v,v'}} \qquad \forall a \in A, v \in V, p \in P \qquad [A8]$$

$$S_{a,v,p} \leq Y_{a,v,v,p+1} \qquad \forall a \in A, v \in V, p \in P \qquad [A9]$$

$$\sum_{a\in A} S_{a,v,p} \leq v\_cap_v \qquad \forall v \in V, p \in P \qquad [A10]$$

$$\sum_{a \in A} (S_{a,v,p} - unavail_{a,v,p}) \geq \sum_{t \in T | V_t, p - t\_periods_t + 1 \leq p^- \leq p} a\_req_t X_{t,v,p^-}$$

$$\forall v \in V, p \in P \qquad [A11]$$

$$S_{a,v,p} - unavail_{a,v,p} \geq \sum_{t \in T | V_t, p - t\_periods_t + 1 \leq p^- \leq p} a\_type\_req_{t,a} X_{t,v,p^-}$$

$$\forall a \in A, v \in V, p \in P \qquad [A12]$$

$$Z \ urs \qquad [A13]$$

$$D_t \geq 0 \qquad \forall t \in T$$

$$LATE_{th} \geq 0 \qquad \forall th \in TH$$

$$PRI\_VIOL_{th,t'} \geq 0 \qquad \forall th \in TH, t' \in T \mid i_t > i_{t'}$$

$$PREC\_VIOL_{t,t'} \geq 0 \qquad \forall t \in T \mid t' \in R_t$$

$$S_{a,v,p} \geq 0 \qquad \forall a \in A, v \in V, p \in P$$

$$W_{a,v,p} \geq 0 \qquad \forall a \in A, v \in V, p \in P$$

$$X_{t,v,p} \in \{0,1\} \qquad \forall t \in T, v \in V_t, p \in P$$

$$Y_{a,v,v',p} \in \{0,1,2,...\} \qquad \forall a \in A, v \in V, v' \in V, p \in P$$

5. Model Description

a. Direct Solution (Monolith)

The objective [A0] assesses a penalty if any “high-priority” test is not finished (high-priority tests must be finished, lowest priority tests can be finished on a not-to-interfere basis), plus a penalty for any test finished after its test priority deadline. There is also a penalty for any higher-priority test that is not started before any lower-priority one is started, and a penalty for any subsequent test begun before a precedent test has been completed. The “cost” units of this objective are “periods.”

Each constraint [A1] requires the objective function value to be at least as high as the time period of the last high-priority test completion.

Each constraint [A2] records whether a test is completed during the planning horizon, or signals that this has not happened.

Each constraint [A3] reckons if some higher-priority test is not completed by its deadline.

Each constraint [A4] records if a higher-priority test is started before a lower-priority one begins, or signals that this has not happened.

Each constraint [A5] records if a precedent test has been completed before a successor begins, or signals that this has not happened.

Each constraint [A6] records if a successor test has begun before a predecessor finishes, or signals that this has not happened.

Each constraint [A7] determines for a newly arrived shipment of test assets with no specific destination where to distribute these among receiving test venues.

Each constraint [A8] accounts for test asset receipts and movements by asset type, facility, and time period.

Each constraint [A9] establishes the number of assets located at a venue and present there during a time period.

Each constraint [A10] limits the number of assets that can be co-located at a test venue during a time period.

Each constraint [A11] allocates available assets to a test venue during a period among various tests plus those assets held from testing for other purposes.

Each constraint [A12] may optionally allocate a particular type of asset to a test venue during a period.

Decision variable domains are given by [A13].

b. Indirect Solution (Cascade)

Instances of this resource-constrained project scheduling integer linear program monolith can be very difficult to solve. In such cases, we might accept a solution with a large integrality gap (the difference between the cost of the solution and a lower bound on how much lower that cost might be for other, as yet undiscovered solutions), and then improve the quality of that solution as will be shown.

We can also employ an indirect method to solve smaller restricted problems in a time cascade. In such a cascade, only decision variables in a restricted window of the planning horizon are free, and all others are fixed. Beginning with a window at the start of the planning horizon, we solve for only the free variables. We then move the window forward in time some number of time periods, fixing behind us the variables leaving the window at their current values. If the window is sufficiently long to include the active tasks and asset movements that are free within it, we achieve a feasible, but sub-optimal solution.

With either a sub-optimal monolith or cascade solution, we can improve solution quality by attempting to move tasks earlier in the time horizon. We fix all decision variables, then, one task at a time, free decision variables for just that task *t*, and attempt to minimize:

$$
\begin{aligned}
MIN \quad & (p + t\_periods_t - 1) X_{t,v,p} \\
& + (\|P\| + 1) \sum_{th \in TH} D_{th} + \sum_{t \in T \setminus TH} D_t + \sum_{th \in TH} penalty_{i_{th}} LATE_{th} \\
& + pri\_pen \sum_{th, \in TH, t' \in T | i_t > i_{t'}} PRI\_VIOL_{th,t'} + prec\_pen \sum_{t \in T | t' \in R_t} PREC\_VIOL_{t,t'} \\
& + move\_pen \sum_{a,v,v' \neq v,p} m\_periods_{v,v'} Y_{a,v,v',p} \qquad \text{[A0P]} \\
& \qquad\qquad X \\
s.t. \quad & \text{[A2-A13],}
\end{aligned}
$$

with all variables fixed except those for test *t*.

This finds gaps in the schedule and slides tasks earlier in the planning horizon while maintaining schedule feasibility, which is desirable. We continue cycling through the tasks, fixing earlier task schedules, until we find no additional improvement.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. MODEL IMPLEMENTATION AND ANALYSIS

A. COMPUTER IMPLEMENTATION

We implement the mathematical model described in Chapter III in the General Algebraic Modeling System (GAMS) (GAMS Development Corporation, 2014) using the IBM ILOG CPLEX mixed-integer programming (MIP) optimization solver. We use a computer with dual 3 GHZ processors, 96 GB RAM and dual 465 GB disk drives. We provide additional functionality within the GAMS script to analyze different aspects of the scheduling problem, which are outlined in this chapter. The GAMS code, parameters and input data for the DT&E project are available from the author, or his advisors. For the following case study, the monolithic optimization model has 39,031 constraints and 29,271 variables, with 10,590 of these discrete variables. Solution times of the monolith can be hours, but we will show how this has been reduced to seconds with a problem cascade.

B. ANALYSIS

1. Case Study Data

This thesis uses test data provided by MARCORSYSCOM from a previously planned DT&E project. The data contains 43 tests, 36 of which are involved in partial orders (i.e., have precedence relationships). All tests are either high or medium priority, and their duration estimates are given in days. There are six available test venues located across the United States, and seven test assets available. There is only one test asset variant, and all test assets are available at the beginning of testing.

2. Methodology and Results

We create schedules and temporal statistics based on the MARCORSYSCOM test data. This data provides the opportunity for one-to-one comparison of model outputs against SME created schedules and DT&E duration estimates. This data also provides a means to verify the model is performing as intended, and validate the model outputs. In this thesis, feasibility refers to a solution that adheres to all inelastic constraints, and identifies any elastic constraint penalties incurred; optimality refers to how close our solution is to a bound on all achievable solutions.

We verify the model is performing as designed by visually inspecting results to ensure that each constraint is satisfied or highlighted otherwise. This is done for test precedence and priority, test and venue pair assignments, and asset availability. Additionally, we verify the test asset movements between venues to ensure proper asset accountability within the network, as well as proper accounting of time required to make the movements.

We validate that the model is providing feasible schedules with deterministic DT&E completion times comparable to those estimated by MARCORSYSCOM (see Table 4). The differences between the MARCORSYSCOM and the GAMS model estimates are within two weeks for three-to-six month schedules. Additionally, the results in Table 4 suggest that current MARCORSYSCOM estimates are overly optimistic in the shortest and longest DT&E completion times. This is most likely due to unintended relaxations of the precedence requirements when heuristically summing individual test durations. The results in Table 4, in conjunction with the verification process described above, suggest that the model is producing valid solutions.

Table 4. Comparative analysis of DT&E duration estimates (days) generated manually by MARCORSYSCOM versus the GAMS model estimates. GAMS model and MARCORSYSCOM estimates are reasonably close (within two weeks for three-to-six month schedules).

Model Validation			
Methods	DT&E Duration (Days)		
	Optimistic	Mean	Pessimisstic
MARCORSYSCOM	66	110	132
GAMS Model	74	100	143

a. Model Tuning

We tune the model by varying penalty parameters, discussed in Chapter III, and assessing any changes in feasibility, optimality, and computation time. We allow test precedence and priority relationships to be either elastic or inelastic constraints.

In practice, the precedence relationships will most likely be treated as inelastic constraints, because many of the predecessor tests ensure basic safety requirements are met before moving on to tests that involve higher risks. However, this is an added restriction that may reveal test-sequencing problems that prevent DT&E completion time by a required date.

As described in Chapter III, we wish to complete higher-priority tasks before lower-priority tasks. However, opportunities exist where there are asset and venue availabilities to complete lower-priority tests, but not higher-priority tests. Additionally, there are instances where a lower-priority test precedes a higher-priority one. In order to take advantage of the available capacity, and adhere to taut precedence constraints, we relax the priority constraints as elastic with no penalty. We give higher-priority tests preference by assigning them earlier deadlines, with higher penalties. For this data, we assign a penalized high-priority test deadline 30 days shorter than the estimated DT&E duration, and a penalized medium-priority test deadline close to the estimated DT&E duration. This improves model solve times and optimality.

Additionally, planners prefer to minimize test asset movements between test venues if possible. We use a small move penalty of 0.1 (asset days per movement day) to minimize unnecessary movements. This makes solutions with more movements less attractive to the solver, speeds up solve times, and eliminates schedule solutions with unnecessary movements.

A thorough study of the DT&E problem described in the previous chapter allows us to eliminate unnecessary elements from the enumerated sets. This is done dynamically in GAMS with the use of conditional statements that limit the number of nodes, and arcs in the modeled network. We make every effort to reduce the dimensionality of the ILP through set enumeration reduction techniques, however the data still presents 39,031 equations, 29,271 single, and 10,590 discrete variables. The reductions in run time and optimality gap (the difference between the value of the best solution found and a bound on how much better a solution might be) can be seen in Table 5. Note the pre-set enumeration reduction results for the optimistic and PERT mean planning estimates. The higher optimality gaps and run times are a function of the longer planning horizons required to complete all tests when considering longer planning estimates. The set enumeration reduction techniques achieve an impressive improvement in both optimality gaps, and run times. From this point forward, we refer to the straightforward solve of the entire ILP network with set enumeration reduction techniques applied as the “monolith.”

Table 5. Results of set enumeration reduction techniques on relative optimality gap (%) and run time (minutes) when solving DT&E schedules with Optimistic and Mean test duration estimates. Set enumeration reduction techniques greatly reduce the optimality gap in much shorter run times for longer planning horizons.

	Optimistic Test Duration Estimates (80 Day Planning Horizon)		PERT MeanTest Duration Estimates (110 Day Planning Horizon)	
	Optimality Gap (%)	Run Time (minutes)	Optimality Gap (%)	Run Time (minutes)
Pre-Set Reduction	22%	34	89%	354
Post-Set Reduction	18%	34	18%	34

As seen in Table 5, solving the monolith with longer planning horizons still has a relatively high computation time for simulation efforts, where we want to run many replications. However, in order to accommodate the entire range of possible test durations we must make the planning horizon at least as long as the pessimistic DT&E duration estimate. Judiciously choosing a planning horizon to be just long enough, along with the use of priority deadlines and associated penalties has coerced the CPLEX solver to find solutions faster. However, these solve times are still not suitable for simulation.

We use the cascade method discussed in Chapter II to further speed up computation time by solving and fixing solutions for smaller ILPs over iteratively larger planning horizons. The cascade method does not necessarily suggest optimal solutions, but does provide feasible, and reasonably optimal solutions (within two weeks for three-to-six month schedules) with much shorter compute times (see Table 6).

Table 6. Comparative analysis of DT&E duration estimates (days) generated by the monolith and cascade solve methods with their associated run times (%). The cascade method drastically reduces model solve times with reasonably optimal solutions (within two weeks for three-to-six month schedules).

Deterministic Differences						
Solve Method	Optimistic Estimates		Mean Estimates		Pessimisstic Estimates	
	DT&E Duration (Days)	Run Time (Minutes)	DT&E Duration (Days)	Run Time (Minutes)	DT&E Duration (Days)	Run Time (Minutes)
Monolith	74	34	100	34	143	34
Cascade	82	2	110	2	158	2

We employ the cascade method with random variables for test durations to simulate overall DT&E durations, which provide a distribution of possible outcomes for temporal statistical analysis (see Table 7 & Figure 2). We generate random test durations using built-in GAMS uniform random number generator and Beta distributions with parameters a and b calculated using Equations 2 and 3.

Table 7. Descriptive statistics for DT&E duration simulation with 200 trials of randomly generated test durations.

Quantiles			Summary Statistics	
100.0%	maximum	138	Mean	117.3
99.5%		138	Std Dev	7.956
97.5%		132	Std Err Mean	0.563
90.0%		127	Upper 95% Mean	118.4
75.0%	quartile	123	Lower 95% Mean	116.2
50.0%	median	118	N	200
25.0%	quartile	112		
10.0%		107		
2.5%		100		
0.5%		98.01		
0.0%	minimum	98		

The mean DT&E duration (117 days) from a 200-trial random test duration simulation is within a week of the cascade model's deterministic result (110 days) using mean test duration estimates (see Tables 6 & 7). The simulation mean is also roughly the same as the mode for the simulation (see Table 7). This may suggest that the MARCORSYSCOM DT&E duration estimate (110 days) developed using the mean test duration estimates is robust for planning and budgeting purposes. However, a closer look at Table 7 reveals that less than 25% of the simulated DT&E durations were 110 days or shorter.

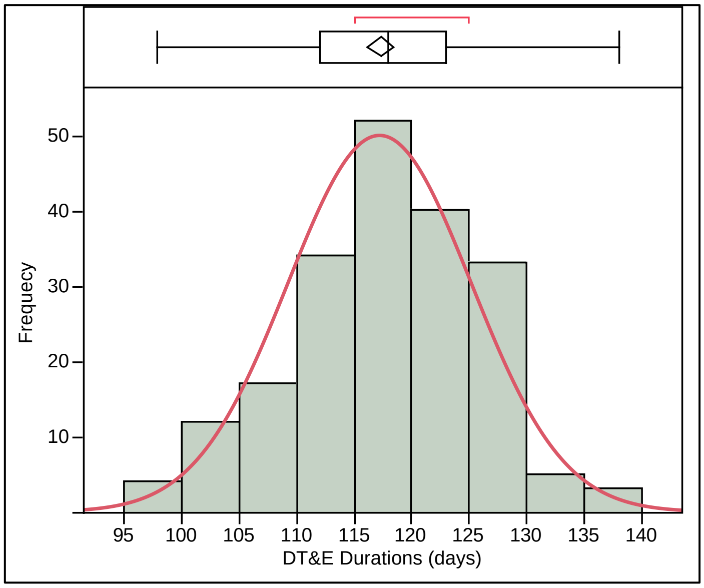

Figure 2. Histogram of DT&E duration (days) simulation with 200 trials of randomly generated test durations. The outcomes appear to be normally distributed.

The full range of simulated DT&E duration outcomes (see Figure 2) provides a better assessment of the temporal risks involved for planning and budgeting purposes than the current MARCORSYSCOM point estimates of optimistic, mean and pessimistic.

In order to make probabilistic statements about these outcomes, we conduct a statistical test to determine if the DT&E duration outcomes can be reasonably represented as normally distributed (see Table 8). The results of a Shapiro-Wilk Test (Shapiro & Wilk, 1965) lead to accepting the null hypothesis that our sample is normally distributed.

Table 8. A Shapiro-Wilk Test leads us to accept the null hypothesis that the results from a 200-trial simulation of DT&E duration can be reasonably represented as normally distributed.

Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.989666	0.1596

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

We also conduct a visual inspection of a Normal quantile-quantile plot (see Figure 3) for further evidence of normality (Wilk & Gnanadesikan, 1968). The Normal probability plot points of the simulated DT&E durations, in black, follow closely with the Normal line (solid red) providing us further confidence to make probabilistic statements about DT&E durations assuming they are normally distributed.

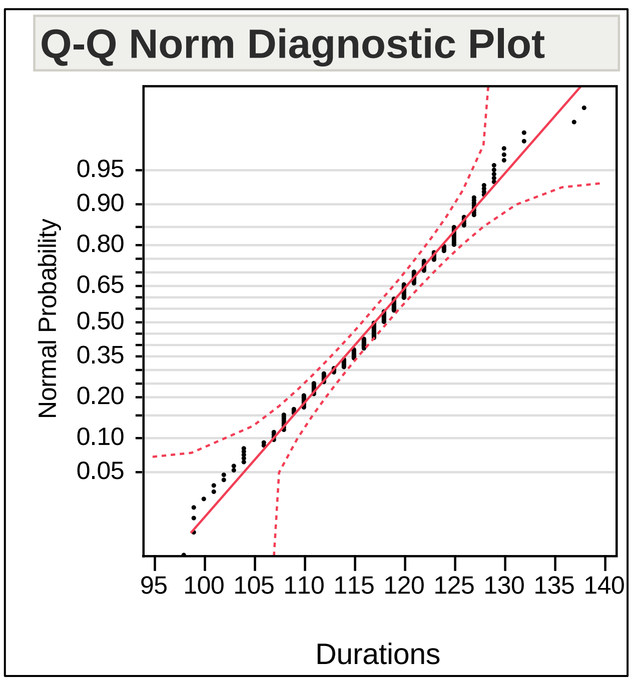

Figure 3. The Quantile-Quantile Normal probability plot points of the simulated DT&E durations (in black) follow closely with the Normal line (solid red) providing us further confidence to make probabilistic statements about DT&E durations assuming they are normally distributed.

By assuming that the DT&E durations are normally distributed, we can make probabilistic statements such as: "what is the probability of completing DT&E within a 110 days (the MARCORSYCOM planning estimate)," or "what is the number of days required to complete DT&E with a probability of 80%?" The answer to the first question is $P(\text{DT\&E Complete} \le 110 \text{ days}) = .18$. This is a fairly low probability of success for making milestone planning and budgeting decisions. However, the answer to the second question $P(\text{DT\&E Complete} \le \text{X days}) = .8$ is 124 days. This may or may not be a big difference (close to three weeks) to the planners and decision makers, but this certainly provides more information for better analysis and decision making than point estimates. This statistical hypothesis test result suggesting a Normal distribution for total test program duration with Beta-distributed test durations is most useful for analysis and inference with this test program, but not for any other test program, or even for any modification of this test program. Case-by-case, test program-by-test program, we must employ similar analyses before assuming normality elsewhere. In fact, it is easy to conjure realistic test programs with Beta-distributed task times that result in far from normally distributed total test program durations.

b. Trade Space Analysis

As discussed in Chapter I, DT&E planners as well as program managers are often concerned with analyzing multiple COAs. We use the mean planning estimates for test durations to assess the trade space between number of available assets and DT&E duration. We use a technique commonly referred to as project crashing, to flood the model with additional assets to determine the shortest amount of time required to finish testing, and then reduce the number of assets available to determine their effect on completion times (see Figure 4). When graphed, this produces a piece-wise linear function that can be visually analyzed for cost-benefit analysis, or to determine the right number of test assets required to complete testing within a given time horizon. This may also be valuable to the DT&E planners to have multiple schedules on hand in the event that more or less assets become available before, or during testing.

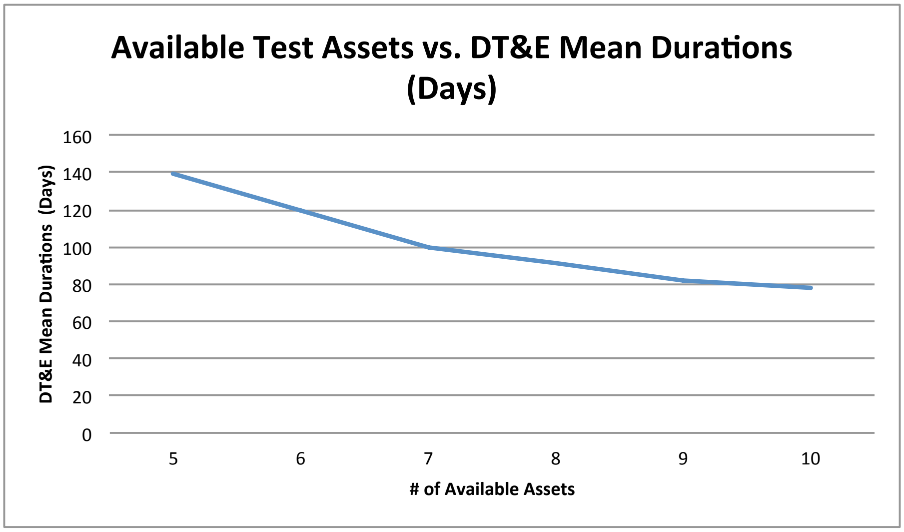

Figure 4. Trade space analysis plot of number of available assets versus DT&E duration (days) using mean test duration estimates. DT&E duration decreases by roughly one month per asset from five-to-seven assets and then two weeks per asset from seven-to-nine assets.

The results in Table 9 provide further evidence that the monolith becomes more difficult to solve (run time) as the resources (test assets) become more limited, and require longer planning horizons. However, as we have shown, this can be ameliorated by indirect cascade solutions.

Table 9. Resulting DT&E durations (days), required planning horizon (days), relative optimality gap (%), and run time (minutes) for given # of available test assets using mean test duration estimates.

Monolith Tradespace Analysis Available Assets vs. DT&E Completion Time (days)				
Available Assets	DT&E Duration (Days)	Planning Horizon (Days)	Relative Optimality Gap (%)	Run Time (Minutes)
5	139	150	34%	34
6	119	130	34%	34
7	100	110	18%	34
8	91	100	13%	34
9	82	90	6%	34
10	78	90	4%	2
20	75	80	0%	1

V. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

This thesis presents an optimization and simulation model as a decision support tool to improve current DT&E scheduling. The proposed model, unlike current manual scheduling techniques, suggests schedules that are feasible, nearly optimal, and are produced in a timely manner for effective analysis of alternatives.

We formulate the difficult and time-consuming DT&E scheduling problem as an integer linear program (ILP). We incorporate set enumeration reduction techniques that reduce monolith solve times while suggesting DT&E schedules and associated duration estimates that are comparable to current methods. The ensured feasibility of these schedules and duration estimates provide additional confidence for coordination, milestone planning and budgeting purposes.

An additional benefit is that should any change occur in scheduling, we can re-schedule quickly, preserving near-term, already-promulgated events, and/or in the longer term, employing methods to reduce turbulence in schedule changes (Brown et al., 1997).

We employ a cascade method to further reduce solve times for simulating DT&E durations with test durations as random variables from the Beta distribution. The simulation results provide a more complete picture of the possible DT&E test duration outcomes for analytic decision making than the current point estimates.

At a minimum, this model can be used to validate the feasibility of manually created schedules.

B. FUTURE WORK

In this study, we provide a model that suggests feasible and conservatively optimal DT&E schedules for increased analytics. Further studies may enhance monolith and cascade solution optimality with reduced run times. Additionally, the data inputs and parameters to this model may be automated from a single user form, possibly a graphical user interface, for ease of use. It is straightforward to embed our model and solver with Microsoft Excel. There is also work that could be done to automate the schedules into a format that is more like the current MARCORSYSCOM final products as well as generate analytical tables and graphs as shown in Chapter IV.

LIST OF REFERENCES

Brown, G., Carlyle, W.M., Harney, R.C., Skroch, E.M., & Wood, R.K. (2009). Interdicting a nuclear weapons project. *Operations Research, 57(4),* 866–877.

Brown, G.G., Dell, R.F., & Wood, R.K. (1997). Optimization and persistence. *Interfaces, 27*, 15–37.

Brown, G. G., Graves, G. W., & Ronen, D. (1987). Scheduling ocean transportation of crude oil. *Management Sciences, 33*(3), 335–346.

Clark, CE. (1962). The PERT model for the distribution of an activity time. *Operations Research, 10*, 405–406.

Davis, R. (2008). Teaching note: Teaching project simulation in Excel using PERT-Beta distributions. *INFORMS Transactions on Education, 8*(3), 139–148.

Defense Acquisition University. (2015) *Defense Acquisition Guidebook*. Web. 18 May 2015.

Demeulemeester, E. L., & Herroelen, W. (2002). *Project scheduling: A research handbook.* Secaucus, NJ, USA: Kluwer Academic Publishers. Retrieved from http://www.ebrary.com

Elmaghraby, S. E. (1977). *Activity networks: Project planning and control by network models.* New York: Wiley.

Fischer, H. (2011). *History of the central limit theorem: From Laplace to Donsker*. London: Springer.

Fulkerson, D.R. (1962). Expected critical path lengths in PERT networks. *Operations Research, 10*, 808–817.

GAMS Development Corporation (2015) GAMS (Version 24.4.1)[Computer software]. Retrieved from http://www.gams.com/

Grose, R. T. (2004). *Cost-constrained project scheduling with task durations and costs that may increase over time: Demonstrated with the U.S. Army future combat systems* (Master's thesis). Retrieved from Calhoun http://calhoun.nps.edu/handle/10945/1582

Kelley, J.E. Jr and M.R. Walker (1959). Critical path planning and scheduling. *Proceedings of the Eastern Joint Computer Conference, 16*, 160–172. doi: 10.1145/1460299.1460318

Kelley, J.E. (1961). Critical path planning and scheduling: mathematical basis. *Operations Research, 9*, 296–320.

Lootsma, F.A. (1966). Network planning with stochastic activity durations: An evaluation of PERT. *Statistica Neerlandica, 20*, 43–69.

Malcolm D.G., Roseboom J.H., Clark C.E., & Fazar, W. (1959). Application of a technique of research and development program evaluation. *Operations Research, 7,* 646–669.

Microsoft. Microsoft Excel (Version 2011) [Computer Software]. Retrieved from www.microsoftstore.com/MicrosoftExcel

Savage, S. (2009). *The flaw of averages: Why we underestimate risk in the face of uncertainty.* Hoboken, N.J.: Wiley.

Shapiro, S.S., Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*(3/4), 591–611.

Trietsch D., Baker K.R. (2012) PERT 21: Fitting PERT/CPM for use in the 21st century. *International Journal of Project Management, 30(4)*, 490–502.

United States Department of Defense. (2015) *Operation of the Defense Acquisition System* (DODI Number 5000.02). Washington, DC: GPO.

Wilk, M.B., Gnanadesikan, R. (1968). Probability plotting methods for the analysis of data. *Biometrika, 55*(1), 1–17.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California